A Supervised Cooperative Learning System for Early Detection of Language Disorders

Martín-Ruiz, M.L.[1], Valero Duboy, M.A.[1], Pau de la Cruz, I.[1], Peñafiel M. [2], Torcal C.[2]

[1]Telematics Systems for the Information Society and Knowledge (T>SIC). Department of Telematic Engineering and Architectures, Universidad Politécnica de Madrid. Spain
[2] Colegio Legamar. Ctra. Leganés-Fuenlabrada Km. 1,5. 28914 Leganés – Madrid
{marisam, mavalero, ipau}@diatel.upm.es
{dirección, infantiluno}@colegiolegamar.es

Abstract. The Quality of Life of a person may depend on early attention to his neurodevelopment disorders in childhood. Identification of language disorders under the age of six years old can speed up required diagnosis and/or treatment processes. This paper details the enhancement of a Clinical Decision Support System (CDSS) aimed to assist pediatricians and language therapists at early identification and referral of language disorders. The system helps to fine tune the Knowledge Base of Language Delays (KBLD) that was already developed and validated in clinical routine with 146 children. Medical experts supported the construction of Gades CDSS by getting scientific consensus from literature and fifteen years of registered use cases of children with language disorders. The current research focuses on an innovative cooperative model that allows the evolution of the KBLD of Gades through the supervised evaluation of the CDSS learnings with experts´ feedback. The deployment of the resulting system is being assessed under a multidisciplinary team of seven experts from the fields of speech therapist, neonatology, pediatrics, and neurology.

Keywords: Information processing for pediatrics health care, Computational support for clinical decisions on language disorders, E-Health applications

1 Introduction

Evidence based research states that neurodevelopment processes from birth until the age of six hardly influence on the Quality of Life of a person. Diagnosis and treatment actions do rely on the detection of developmental disorders in early childhood. The prevalence of language delays is the second highest within the common developmental disorders (1-19%) [1], and it is often associated with negative long-term outcomes [2,3,4,5]. Accurate diagnosis of children's health status leads to better development results. Unfortunately, pediatricians and educators at nursery school do not always detect the existence of language disorders at early stages. This evidence demands to develop medical information systems to facilitate non delayed referrals to the specialists.

Thus, experts could identify disorders in the first neurodevelopment stages, suggest medical treatments and share them with authorized professionals involved in the child's evolution. Gades Clinical Decision Support System (CDSS) and a Knowledge Base of Language Delays (KBLD) were developed in 2011-12 in order to provide pre-school teachers, speech therapists and parents with monitoring tools for children with language disorders [6]. The deployment and validation of these solutions showed an 84% success rate when evaluated, along three months, with 146 children who attend to Legamar (Madrid) nursery school. A systematic review of the field trials identified the necessity to create a model ready to optimize the KBLD starting from the erroneous suggested decisions, ambiguous questions and experts' feedback.

This paper details the analysis and design of a system that facilitates a supervised cooperative learning process for Gades CDSS, which is very important in order to obtain the correct evolution of the CDSS. The resulting model contributes to enhance the KBLD and the success rate of early referral of children with language disorders at the clinical and educational scenarios foreseen at this research program.

2 Background

The positive impact of intervention at the transitory nature of a certain development disorder was pointed out in 2005 in the white paper on early attention edited by the National Foundation for Prevention and Care of People with Disabilities [7]. These disorders may become a significant deviation from the expected development and result from health or relational events that compromise biological, psychological and social development. Consequently, early attention focuses on the detection of possible alterations in child development to trigger the activation of requested procedures [6].

PubMed, JAMIA and ModernMedicine are valuable sources to make a review about the published systems that improve early attention on children. The systematic analysis that was done for this research looked up the following three keywords: *early attention*, *language disorders* and *e-Health* [8,9,10,11,12]. As a result, no solutions were found about the utilization of CDSS to assist teachers and language therapists in the early detection of language disorders. Although most of identified e-health solutions deal with the management of the Electronic Medical Record (EMR) and its interoperability barriers, none of them do provide the application of Knowledge Based Systems and EHRs information to suggest early referral processes in pediatric care.

The potential benefits of CDSSs in medicine were identified since the early 70s and received much attention over the past three decades. Expectations were raised about its potential to increase productivity in a medical environment, to support diagnoses and other types of medical decisions, and to assist in medical training [13]. Moreover, CDSS definitely enhances its impact on health care outcomes as well as it is hardly connected with Electronic Medical Records (EMR) at the point of care.

The benefits of a CDSS to increase quality and patient safety, improve the patients adherence to guidelines for prevention and treatment, and to avoid medication errors was related by recent studies [14, 15]. CDSS have been typically validated in the domains of diabetes mellitus, coronary artery disease and hypertension but not to prevent children disorders such us language or cognition delays.

Children usually acquire good verbal communication by the age of three years. However, current rates of detection of development disorders are lower than their real incidence [16], which means that early identification of children with such disorders remains a pending task. Thus, Gades CDSS was deployed to provide pre-school teachers and speech therapists with a detection tool of language disorders for children.

The Office of Special Education and Student Services of the Virginia Department of Education states that: "the development of communication skills is important for all students and can impact school success" [17]. In the same way, pre-school teachers have not specific formation about child development disorders. Therefore pre-school teachers could support the early detection of future development pathologies by using KBLD tools. Hence, the deployment of the Gades CDSS at a nursery school becomes of greater interest as detailed in previous works [6], [18].

Gades helps to detect neurological disorders in children in order to prevent added pathologies, achieve functional improvements and allow a more adaptive adjustment between a child and his or her surroundings. Gades does not include any dynamic learning mechanism. This paper details an autonomous solution to provide Gades CDSS with and effective supervised cooperative learning mechanism.

3 Methods

Current research starts from Artificial Intelligent (AI) techniques whose aim is to develop systems and agents that learn automatically. AI identifies three main learning types, according to the feedback provided by the learning agent [19]:

- "In unsupervised learning, the agent learn patterns in the input even though no explicit feedback is supplied"
- "In reinforcement learning, the agent learn from a series of reinforcements – rewards or punishments"
- "In supervised learning, the agent analyzes some example input-output pairs and learns a function that maps from input to output"

A comparative review of these technologies showed that none of them were ready to solve Gades CDSS cooperative learning requirement, since these technologies infer new knowledge from existing knowledge. KBLD evolution process does not require generating new knowledge as experts have this knowledge. The new system must focus on knowledge agreements coming from the experts' experience. Hence, the KB should dynamically evolve to achieve effective Gades performance.

The resulting enhancement model needs supervised cooperative processes so that the correct refinement of the KB may fit to the whole life cycle of Gades according to the diverse use cases of children with language disorders.

The deployment of the proposed system needs a distributed application architecture to permit concurrent inputs from the involved specialists. A web platform, linked to

the available Gades solution, will facilitate the KBLD evolution to take into account the experts' knowledge and experience.
The enhancement of the KBLD relies on a supervised cooperative learning model to be carried out by the neuropediatricians, psychologists and language therapists, or other experts in charge of the evaluation of the language acquisition level of the child. These users use the Gades CDSS and detect refinement needs in their daily routine.

The methodology defined for this research includes two incremental phases:

1. Language evaluation phase: This phase encompasses the results gathered from the Language Evaluation (LE) of children enrolled at a nursery school by means of the observation performed by the pre-school teachers and the cooperative work of the nursery school language therapist (see steps 1, 2, 3, 4 and 5 at Fig. 1).

2. Supervised cooperative learning: The language therapist can provide suggestions for Knowledge Evolution (KE) along the LE stage. These proposals may enrich an evolved KBLD coming from the inputs received from the language therapist and the team of experts. The supervision staff is composed by a neuropediatrician, two language therapists, and a neonatologist who assessed the construction process of Gades CDSS.

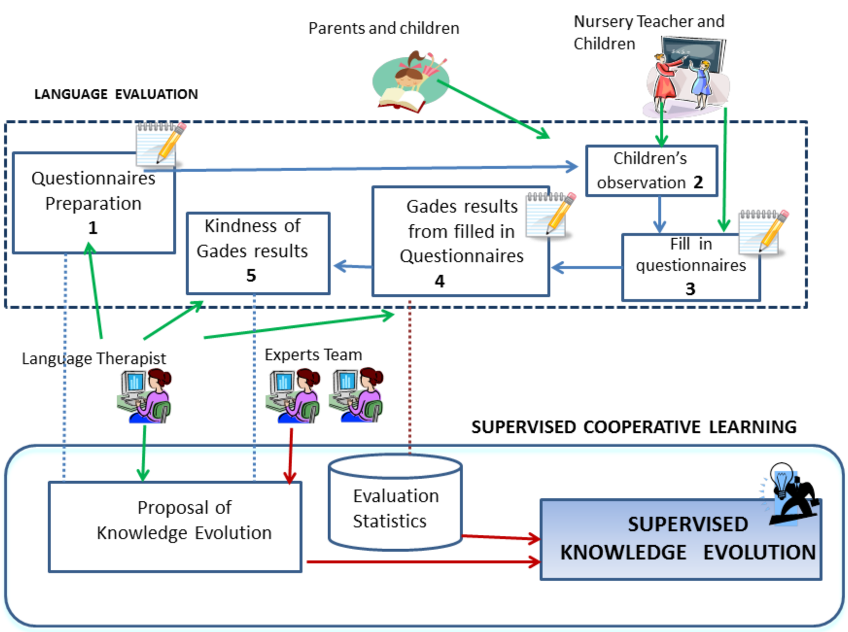


Fig. 1. System Methodology stages

The KE process stage leads to the automatically generation of an updated OWL (Ontology Web Language) file that contains the KBLD [6] without the intervention of an engineer. The richness of OWL semantic technologies makes easier to express the KBLD and contributes to the interoperability of the deployed E-Health solution.

4 Results

This section details the main results of this research: 1) the analysis of the system requirements and UML (Unified Modelling Language) use case diagrams for Gades CDSS supervised learning process; 2) the designed architecture of the system; and 3) the deployment of the final solution that allows the Gades supervised cooperative learning service as used at real scenarios like the nursery schools.

4.1 System Analysis

Once the GADES knowledge evolution system objectives have been defined we proceeded to the analysis of its functionalities through a formal definition of requirements.

The requirements have been obtained by the discussion of several use cases by a multidisciplinary team (the same as used for the GADES development [18]). The functional requirements obtained through the discussion have been refined by an iterative process.

Table 1 summarize the requirements obtained grouped in three different areas. In the table are shown the general overviews of the requirements that have guided the development of the system.

Table 1. System update functional requirements

Knowledge evolution principles	The evolution of knowledge will be guided by authoritative experts in all cases. Although there is the possibility that the system includes requests automatically, the team of experts will take the final decision.
	The system must facilitate the flow of knowledge among experts fostering a constructive discussion and providing consensus mechanism for final decisions.
	The system will allow refinement of existing knowledge by assigning weights according to the importance of the questions. The inclusion of a request for refinement in the system can be carried out manually by any user with the capability, or automatically upon detection of a pattern by the system.
	The system should allow the inclusion of new knowledge that is not based on the previous one (not a refinement process). In this case the system cannot make requests automatically.
	A set of profiles must be defined to differentiate the user roles. The system must be flexible enough to include new roles with new capabilities at any time.

	The work done by the users of the application (experts) must be accounted. The system must be able to create reports describing the contributions of each of the experts in the evolution of knowledge.
Ontology management	Replication techniques will be designed to avoid the accidental loss of the ontologies physical containers (commonly computer files) since the ontologies formalize the existing knowledge in the solution used.
	The ontologies will be secured by techniques of confidentiality, integrity and control access to ensure their validity.
	The ontology must be controlled with a versioning scheme. This way the system can know the version of the ontology that has been used for a specific decision.
Knowledge evolution system management	The system must allow the integration with different authentication and authorization platforms existing at present. The identification must be personal so it can be multiple identities within the same organization.
	The interaction subsystem will be dependent on the profile and personal preferences of the user accessing the system. The criteria for the interaction design will be the usability and efficiency in obtaining informed consensus of experts.
	All the actions on the system will be conveniently recorded. Registration must allow the reconstruction of the activities held by both the system and each of its members.

Fig. 2 illustrates the UML use case diagram of the Gades Knowledge Evolution Model (GKEM). The actors who interact with the system are *experts in language delays* and *language therapists* at the school. The tasks associated to the use case begin with the language evaluation performed by the language therapist through the observation work carried out by nursery pre-school teachers. Experience acquired along the last year shows that this process is more reliable if started in the last quarter of the year, since teachers have more information about the language acquisition level of their students.

The Use case *Proposal of Gades Knowledge Evolution* starts when a language therapist fills the questions about the language abilities of a child. Next, she/he may decide the questions of the evaluation process to be refined. Users may consider that a question should be raised at a different time and propose to modify it, to remove it or to enrich with new questions.

Use case *Evaluation Statistics Results* extends the collection of answers gathered at the learning evolution process in order to generate statistical results that provide, as an example, the frequency of affirmative or negative answers of individualized items. Whether a scheduled process discovers affirmative or negative result in 100% of evaluations of an item, a refinement is triggered to delete or modify the content of the question.

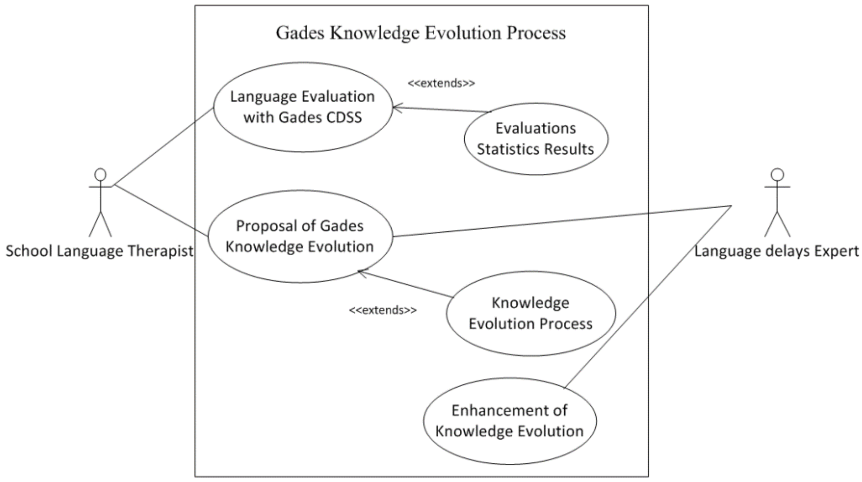


Fig. 2. Gades Knowledge Evolution Process

The use case *Enhancement of Knowledge Evolution* provides the language delays expert with the facility to suggest alternatives to the KBLD by adding these decisions and contribute to the GKEM. A cooperative agreement system proposes experts to reach a consensus about how to include the suggested KB refinement.

Fig. 3 shows the UML sequence diagram of the use case Proposal for Gades Knowledge Evolution. DAO component allow information updates in the database.

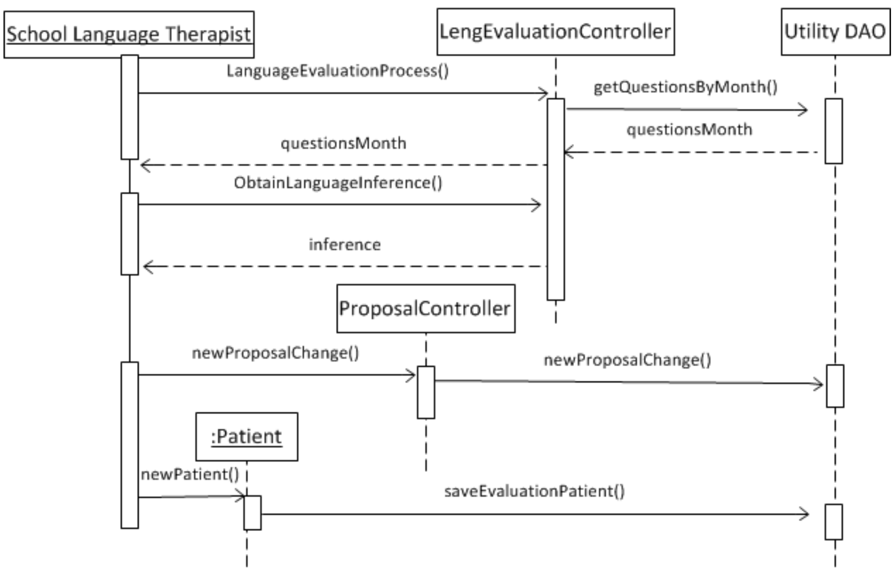


Fig. 3. Use case for Proposal of Gades Knowledge Evolution

4.2 General System Architecture

The functional architecture of the proposed system enhances the interaction among the actors involved, the information management platforms, the reasoning models and the processes that run at each health care scenario. Fig. 4 enumerates the dynamic interaction by paying attention to the interoperation steps.

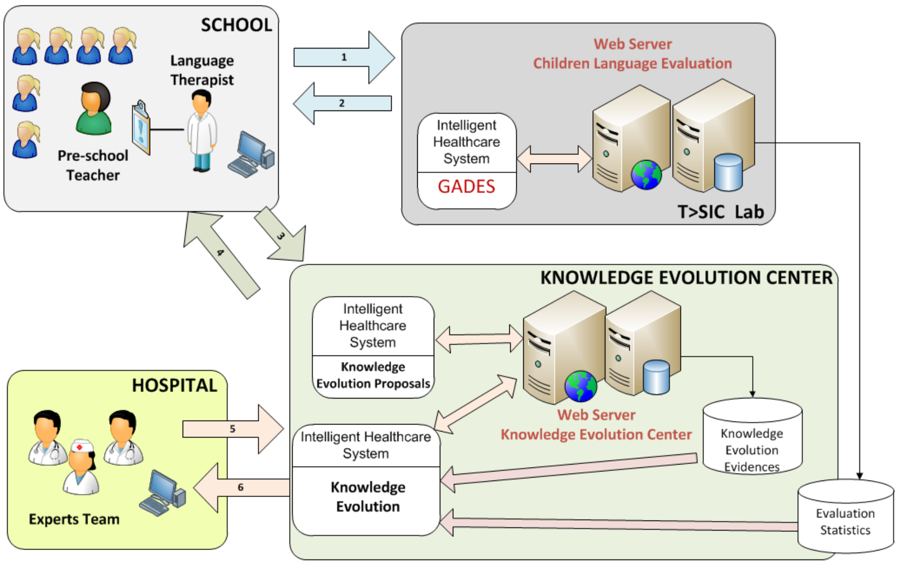


Fig. 4. Global System Architecture

1. Once the process of language acquisition observation concludes at the school, the language therapist decides to use the Gades CDSS in order to assess whether there is a language disorder in the children.
2. The CDSS returns the outcomes to the language therapist. Besides, statistical data is generated from the evaluation stage to complement the KE process. The following steps could happen: A) the result suggested is normal development of the child, B) the result proposed is to derive from the relevant specialist hospital, or C) the system suggests to reevaluate the language acquisition of the child in one, two or three months. Furthermore, the language therapist can determinate if there is a modification proposal for the KB.
3 and 4. The language therapist carries out the KE process. The outcome of this stage is the Knowledge Evolution Evidences.
5 and 6. Experts review the KE proposals and trigger KE process supported by the KE evidences and evaluation statistics.

4.3 Deployment of the KBLD Evolution

This section details the deployment diagram as structured according to the necessary hardware nodes and the relationship between components that enable the evolution of KBLD that allow Gades supervised cooperative learning. Fig. 5 shows the connection of the nodes involved in the execution stage, and the embedded components.

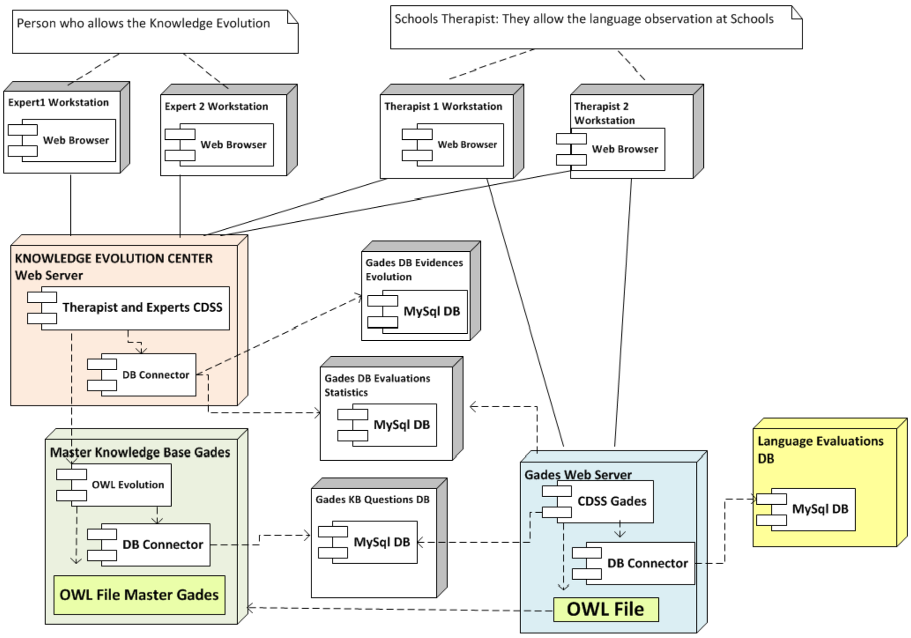
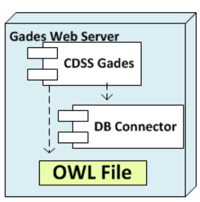
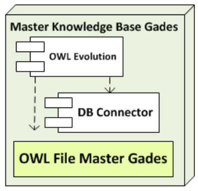
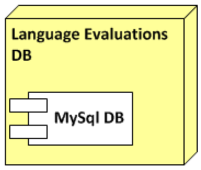


Fig. 5. Deployment of the KBLD Evolution

In this scenario, the KBLD evolution is based on the school observation of language acquisition level in children. Main nodes deployed are (see Table 2):

Table 2. Main nodes for the school observation scenario

Node	Description
	**Gades Web Server**: It brings together some resources entrusted to carry out the learning evolution process. It is composed by the DB connection and the Gades Clinical Decision Support System. This component uses the last version of the OWL File to obtain inferences.

Node	Description
	**Master Knowledge Base Gades**. It carries out the evolution process of Gades CDSS. Gades development and validation is detailed in previous works [6], [18].
	**Language Evaluations Data Base.** It is a component that collects critical information. It is the only place where patient information is stored and requires a secure storage process in the database.

The evaluation process generates evaluations statistics that calculate the number of questions with `yes´ or `no´ answer. This result allows proposing automatic decisions for removal or modification of some KBLD questions. In the same way, the Gades KB questions are stored in a database.

It is planned to obtain the supervised learning evolution with the involvement of primary care paediatricians in the observation process. Hence, the primary care centre becomes a second experimental scenario. A new system is under deployment to use available KBLD and provide primary care paediatricians with the facility to assess the level of language acquisition of their children. The observation process in this scenario is different compared to Gades since the KE is based on two or more observation stages. The achievement of the new KBLD at this scenario will require a KE Centre Web Server. This node appears to centralize the process of learning evolution. Thus, the mission of this node is to centralize all required tasks, necessary to achieve evolution, whenever more information from a language observation is required. This node provides web apps that allow the development of knowledge, both for experts and speech therapists.

5 Conclusions

Validation stage showed that open web applications allowed the easy utilization of the KB by the language therapist. The developed system provide professionals with a usable CDSS for early attention in school as verified with 146 children.

The involvement of 4 experts in neuropediatrics, neonatology and language disorders was crucial for both defining the problem and for refining the KB with existing experience. Gades CDSS identified 7 children with a language delay that had not been previously detected by the school therapist or educator. The alarms generated by the system triggered an evidence based diagnosis for this 4.8% ratio of undetected cases.

The process of Supervised Cooperative Learning of the KBLD is required to achieve later language evaluations at daily routine with the therapist and educators.

Experts agree that the system offers an effective and innovative solution for the complex task of Supervised Cooperative Learning for a KBLD in order to enhance the early detection of language disorders with children aged 0 to 6 years old.

6 Acknowledgments

Mrs. Paloma Tejeda, at Language Intervention Center (LIC) La Salle Campus Madrid, Universidad Autónoma de Madrid.
José Arizcun, MD, neonatologist in developmental disorders and child disability.
Beatriz Chiclana, MD, and Erwin Kirchschlager, MD, pediatricians at Jazmin Public Health Center. María Teresa Ferrando, MD, neuropediatrician at Quiron Hospital of Madrid.

This article is part of research conducted under Talisec+ project (A Framework for knowledge-based management of accessible security guarantees for personal autonomy; TIN2010-20510-C04-01), supported by the Ministry of Education and Science of Spain through the National Plan for R+D+I (research, development, and innovation).

7 References

1. Law, J., Boyle, J., Harris, F., Harkness, A., Nye, C. Screening for speech and language delay: a systematic review of the literature. Health. Technol. Assess. 2, 1–184 (1998)
2. Wilson, P., McQuaige, F., Thompson, L., McConnachie, A. Language Delay Is Not Predictable from Available Risk Factors. ScientificWorldJournal (2013)
3. Law, J. Short- and Long-Term Outcomes for children with Primary Language Impairment (PLI). The Encyclopedia of Language and Literacy Development. The University of Western Ontario (2009)
4. Tomblin, J.B., Records, N.L., Buckwalter, P., Zhang, X., Smith, E., O'Brien, M. Prevalence of specific language impairment in kindergarten children. J. Speech. Lang. Hear. Res. 40, 1245–1260 (1997)
5. Gillberg, C. The ESSENCE in child psychiatry: early symptomatic syndromes eliciting neurodevelopmental clinical examinations. Res. Dev. Disabil. 31, 1543–1552. (2010)
6. Martin-Ruiz, M.L., Valero Duboy, M.A., Pau de la Cruz, I., Ferrando Lucas, M.T., Penafiel Puerto, M. Development of a Knowledge Base for smart screening of language disorders in primary care. 12th International Conference on Bioinformatics & Bioengineering (BIBE), pp. 121–126. IEEE (2012)
7. Lyman, J.A., Cohn, W.F., Bloomrosen M., Detmer D.E. Clinical decision support: progress and opportunities. J. Am. Med. Inform. Assoc. 17, 487–492 (2010)
8. Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., Kotz, S.A., Tillmann, B. Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. Neuropsychology. 27, 121–131 (2013)
9. Naddy, C. The impact of paediatric early warning systems. Nurs. Child. Young. People. 24, 14–17 (2012)
10. McLellan, M.C., Connor, J.A. The cardiac children's hospital early warning score (C-CHEWS). J. Pediatr. Nurs. 28, 171–178 (2013)
11. Parshuram, C., Hutchison, J., Middaugh, K. Development and initial validation of the bedside paediatric early warning system score. Critical Care (2009)

12. Haines, C., Perrott, M., Weir, P. Promoting care for acutely ill children-development and evaluation of a paediatric early warning tool. Intensive. Crit. Care. Nurs. 22, 73–81 (2006).
13. Tolmie CJ, du Plessis JP. The use of knowledge-based systems in medicine in developing countries: a luxury or a necessity?. Methods. Inf. Med. 36, 154–163 (1997)
14. Middleton B. The clinical decision support consortium. Stud. Health. Technol. Inform. 150, 26–30 (2009)
15. Grupo de atención temprana. Libro blanco de la atención temprana. Real patronato de prevención y de atención a personas con minusvalía, 3rd ed., 2005.
16. Council on Children With Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee and Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home. Pediatrics. 118, 405–420 (2006)
17. Speech-Language Pathology Services in Schools: Guidelines for Best Practice. Virginia Department of Education Division of Special Education and Student Services, 2011.
18. Martin-Ruiz, M.L., Valero Duboy, M.A., Pau de la Cruz, I. Deployment and Validation of a Smart System for Screening of Language Disorders in Primary Care. Sensors. 13, 7522-7545 (2013)
19. Russell, S.J., Norvig, P. Artificial Intelligence A Modern Approach. Third Edition. Pearson. (2010)